

Rec'd PCT/PTO **6 APR 2005**
VLIW PROCESSOR WITH POWER SAVING

The invention relates to a data processing apparatus, such as a VLIW (Very Long Instruction Word) processor, that is capable of executing a plurality of instructions from an instruction word in parallel.

5 A VLIW processor makes it possible to execute programs with a high degree of instruction parallelism. Conventionally, in each instruction cycle the VLIW processor fetches an instruction word that contains a fixed number, greater than one, of instructions (often called operations). The VLIW processor executes these operations in parallel in the same instruction cycle (or cycles). For this purpose the VLIW processor contains a plurality of functional units, each capable of executing one of the operations from the instruction word
10 at a time. Different kinds of functional units are typically provided, such as ALU's (arithmetic logics units), multipliers, branch control units, memory access units etc. Often dedicated purpose functional units are also included, designed to speed up programs for a particular applications. Thus, for example, functional units for performing parts of MPEG encoding or decoding may be added.

15 In advanced VLIW processors hundreds of functional units may be present. In principle, the instruction word may contain instructions for all of these functional units in parallel. Often the functional units are organized into groups of one or more functional unit, an instruction word providing one instruction per group. When at least some of the groups contain more than one functional unit grouping limits the length of the instruction word,
20 without reducing the number of functional units.

 All functional units inevitably consume power supply current. When a VLIW processor contains many functional units that operate in parallel, therefore, considerable power consumption occurs. This is inconsistent with requirements for battery-operated apparatuses. It may also increase the cost of cooling measures needed to operate the VLIW
25 processor in a single package, due to the heating associated with power consumption.

 US patent No. 5,815,725 describes the use of clock gating to reduce power consumption in a microprocessor. A monitor circuit monitors whether the microprocessor enters a low activity operational state and if so it gates clock signals to the microprocessor. In US 5,815,725 the clock gating involves disabling the clock signal in only part of the clock

cycles, because the microprocessor must continue to operate. US patent No 5,661,751 describes clock gating during which the clock signal to a peripheral of a microprocessor (a UART) is completely disabled. Similarly, US patent No. 6,345,336 describes disabling of clock signals of part of a cache memory.

5 Clock gating reduces power consumption but when applied to the instruction execution part of a processor it has the disadvantage that it reduces the capability of executing instructions. Significantly, US 5,661,751 and US 6,345,336 apply clock gating to peripheral or auxiliary circuits and not to the instruction execution circuit or the whole instruction memory. US 5,815,725 attempts to mitigate the problem of complete disabling of
10 the clock signal of the microprocessor by disabling the clock signal only in part of the clock cycles. Nevertheless the rate of instruction execution is reduced.

 Among others, it is an object of the invention to provide for a data processing apparatus which uses power saving measures during execution of instructions to reduce power supply consumption without reducing the rate at which instructions can be executed.

15 The invention provides for a data processing apparatus according to Claim 1. This data processing apparatus is of a type, such as a VLIW processor, that processes instruction words that each contain a plurality of instructions. Different functional units execute the instructions from an instruction word in parallel. According to the invention the processing apparatus is constructed so that it is made possible to apply power saving
20 measures, such as clock gating, selectively to part of the functional units and/or memory units that supply instructions to respective ones of the functional units or groups of functional units, dependent on program execution. In the memory units in particular much power can be saved.

 The invention is based on the insight that there exist useful application
25 programs in which the utilization of functional units varies from one program section to another. In such applications it can be determined in advance which functional units will be used in which section. For example, in a program that involves MPEG encoding, specialized functional units for specific tasks in such encoding are only used in specific sections. When the processor executes instruction words from a program section power saving may be used
30 to disable clock signals of the functional units and/or memory units that are known not to be used in that section.

 When the instruction word contains a field dedicated for instructions for a functional unit in which power saving measures are applied, the apparatus may automatically also apply power saving measures to the section of instruction memory that provides that

field when clock gating is applied to that functional unit. More generally, the processing apparatus may apply power saving measures to any resources, such as a register file or peripheral circuits, that are dedicated to the functional unit to which power saving measures are applied.

5 It has also been discovered that in many useful application programs the utilization of different functional units is correlated. In a program section where one functional unit is not used, certain correlated functional units are not used either. Therefore it is advantageous to combine such functional units into a group and to arrange the circuit so that clock gating disables clock signals to all functional units in the group. When the group
10 contains no functional units that are used in a program section, clock gating can be used to disable clock signals to all of the functional units of the group. Moreover, when resources are shared per group of processors clock gating can also be applied to the resources.

15 These and other objects and advantageous aspects of the processing apparatus and method of processing according to the invention will be described in more detail with reference to Figure 1, which shows a processing apparatus.

20 Figure 1 shows a processing apparatus that contains a memory system 10, with memory units 12a-g, a controller 14, and an instruction execution unit 7 that contains groups 70a-g of functional units 18a-c, a register file 72 and an instruction address counter unit 74. Instruction address counter unit 74 has an instruction address output coupled to controller 14. Controller 14 has selection outputs 16 coupled to memory units 12a-g and to groups 70a-g.
25 Furthermore, controller 14 has address outputs coupled to memory units 12a-g. Memory units 12a-g have instruction outputs coupled to respective ones of groups 70a-g and to register file 72. Register file has operand/result output/input ports (not shown separately) coupled to groups 70a-g. Groups 70a-g each contain one or more functional unit 18a-c (the functional units of only one group being shown explicitly), which all have clock gating inputs coupled
30 to the selection outputs 16 of controller 14, operation code inputs coupled to memory units 12a-g, operand inputs coupled to register file 72 and result outputs coupled to register file 72 (all except the clock gating inputs being symbolized by a single connection between memory units 12a-g, groups 70a-g of functional units 18a-c and the register file 72.). One of groups 70a-g has a branch address output coupled to instruction address counter unit 74.

In operation the processing apparatus operates in successive instruction cycles. In successive instruction cycles address counter unit 74 outputs addresses of successive instructions to controller 14 (these instructions will be called "successive" because the corresponding instructions are executed successively, although in the case of branches the addresses may not be successive). Controller outputs further instruction addresses derived from the instruction address to memory units 12a-g. The further instruction addresses address instruction memory locations in memory units 12a-g. Memory units 12a-g output instructions from the addresses to instruction execution unit 7. The combination of instructions output from memory units 12a-g forms an instruction word with fields for the various instructions.

Controller 14 also outputs selection signals which are applied to the memory units 12a-g. Each selection signal indicates whether an instruction from a respective memory unit 12a-g is needed for the current instruction cycle. When the selection signal indicates that no instruction is needed from a particular memory unit 12a-g the memory unit is switched to a power saving state, for example by disabling clock signals in that particular memory unit 12a-g. These clock signals include for example the clock signal that signals the output driver of the memory unit to change the instruction output from the particular memory unit 12a-g, or the clock signal used to precharge bit lines and/or word lines etc. When these clock signals are disabled power is saved, for example because no charging current for outputs, bit lines and/or word lines is needed. Other ways of saving power include disconnecting a power supply source from circuits that need not retain a state during power saving.

Each group 70a-g of functional units 18a-c receives an instruction from a respective one of memory units 12a-g and the selection signal that is applied to that memory unit 12a-g. The selection signal controls whether the group of functional units is switched to a power saving state, for example by disabling clock signals in the functional units 18a-c in groups 70a-g. The disabled clock signals include for example clock signals that cause logic transitions in the output signals from output drivers of functional units 18a-c, or clock signals involved in precharging signal lines. Also, some functional unit contain data memory that consumes less power when the clock is disabled. When these clock signals are disabled power is saved, for example because no charge current for outputs or signal lines is needed.

In those groups where the selection signal does not indicate that clock signals should be disabled, the functional units 18a-c of the group 70a-g determine which of the functional units 18a-c of the group 70a-g should execute the instruction from the corresponding memory unit 12a-g, and that functional unit reads operands addressed by the instruction from register file 72 (if any) and supplies results to register file 72 (if any).

Although it is preferred that clocks are disabled both in cooperating memory units 12a-g and in groups 70a-g, it will be understood that a power advantage is already gained when clock signals are disabled in only one of them.

Controller 14 is capable of selecting and deselecting memory units 12a-g and/or groups 70a-g independently of other memory units 12a-g and/or groups 70a-g. Selection may be controlled in various ways. In one embodiment memory mapping information is used that is loaded into a control memory (not shown) in controller 14 prior to execution of a program of instruction words from memory units 12a-g. In this case the memory mapping information indicates for a number of address ranges of instruction addresses from instruction address counter unit 74 which of the selection signals should be activated. When controller 14 receives an instruction address from instruction address counter unit 74 it detects the address range that contains the instruction address and supplies the selection signals stored for that address.

In another embodiment subsequent switching off or on of selected selection signals is commanded from the instruction words that are executed by execution unit 7. For this purpose a special selection control functional unit may be provided in one of the groups 70a-c, that executes instructions which contain indications of the groups 70a-g that should receive selection signals. Such an instruction may for example be in the form of a mask with respective bits for different groups, to indicate whether or not the group should be selected or not, or in the form of numbers that indicate a group whose selection should be activated or deactivated. Thus, different subsets of (groups of) functional units in which clock signals are to be disabled can be selected. In an extremely simple embodiment, wherein clock signals can be disabled only in one such subset, the command need not specify the subset.

Although figure 1 shows that all groups 70a-g receive selection signals, it will be understood that the invention is not limited to use of selection signals for all groups. In practice controller 14 may not have a selection output for some of the groups 70a-g and these some of the groups may not have a selection input. Thus, these groups are always active. Preferably, at least one group is always active. Also, although each group is shown to receive its own independently settable selection signal, it will be understood that in practice some groups may receive a shared selection signal. Furthermore, although all groups have been shown without distinctions, it will be understood that the groups may in fact differ: functional units in some groups may receive literal data, such as branch addresses or constants from memory units 12a-g, whereas others merely receive operation codes, data being supplied

from register file 72, some groups may receive larger numbers of operands than others, or produce larger numbers of results.

As shown, one of the groups 70a-g has a connection from a branch functional unit (not shown) to update the instruction address in instruction address counter unit 74 in response to an instruction. The branch functional unit executes this update for example when it determines that some condition has been met. Updates may be absolute (replacement of program counter value in address counter unit 74) or relative (addition to the program counter value). This is shown by way of example. In practice more than one group 70a-g may contain one or more branch functional units coupled to instruction address counter unit 74.

Furthermore, although separate memory units 12a-g have been shown for respective groups of functional units 70a-g, it will be understood that some groups may share a memory unit 12a-g, so that the memory unit produces instructions for these groups in parallel (in general these memory units will have wider instruction output than other ones of memory units 12a-g). Of course, clock signals are disabled in such a memory unit, if at all, only when none of the groups of functional units 70a-g that is connected to the memory units 12a-g needs an instruction. This can be implemented using a detector to determine whether none of the relevant groups of functional units needs an instruction, or it may be indicated by instructions from the program.

Furthermore, in some designs register file 72 may be split into a number of register files, some of which are coupled only to a subset of groups 70a-g of functional units 18a-c, sometimes even only to one group 70a-g, in which case that register file can be regarded as part of the relevant group. In the latter case, power saving may be applied to the register file that is only connected to one of the groups 70a-g that is not currently selected, for example by disabling clock signals in that register file. When more than one group has access to a register file power saving may be applied to that register file when the selection signals from controller 14 disable clock signals in all of the groups that have access to the register file. Controller 14 may be provided with a separate selection output for this register file for this purpose, so that power saving in the register file can explicitly be controlled. Alternatively, a detection circuit may be provided to detect whether the selection signals of all involved groups 70a-g signal that power saving should be applied and if so the detection circuit signals that power saving should be applied to the register file as well.

In practice the processing apparatus may use pipelining of instruction execution. That is, in the same instruction cycle controller 14 may process one instruction address, memory units 12a-g may retrieve instructions for a preceding instruction address and

functional units 18a-c may process one or more processing stages for one or more yet further preceding instruction address. In this case, power saving or more particularly disabling of clock signals may also be pipelined, for example by delaying the selection signals from controller 14 by different numbers of instruction cycles for memory units 12a-g and different pipeline stages of functional units 18a-c.

Prior to program execution, it should preferably be determined which program parts need which groups 70a-g of functional units. This is a matter of taking account of specialized functions of the functional units 18a-c, but it may also depend on the different required amounts of parallelism in different parts of the program. For example, a higher amount of parallelism may be needed inside an inner loop.

Programming of the data processing apparatus starts with determination of a description of the operations that have to be performed, for example compiled from a program in a high level computer language. Subsequently, a step is performed to map the operations to functional units. This mapping step allows for some mapping freedom. For example, some arithmetic and logic operations could be performed sequentially on one arithmetic logic functional unit, or in parallel on different arithmetic logic functional units. During the mapping step, an inner loop and surrounding parts of the program may be identified (which are executed many times and only once or a few times, respectively, each time when the program is executed). In this case, the operations of the inner loop are preferably mapped to allow parallel execution in different functional units, whereas operations from the surrounding parts are preferably mapped to one or a limited subset of the functional units, using sequential execution. Moreover, during the mapping step some operations can only be mapped to specific functional units or a group of functional units. Certain MPEG encoding or encoding functions are examples of this.

In a selection step, the combinations of (groups of) functional units that are used in respective sections of the program are identified and information is compiled that indicates which combinations are used in which sections. This information is subsequently used during execution of the program to disable clock signals selectively in those (groups of) functional units that are not used in a section when instructions from the section are executed, for example in the form of memory mapping information or in the form of commands to disable or enable clock signals in selected functional units.